MATH 556 - PRACTICE EXAM QUESTIONS

1. The joint pdf for continuous random variables X, Y with ranges $\mathbb{X} \equiv \mathbb{Y} \equiv R^+$ is given by

$$f_{X,Y}(x,y) = c_1 \exp\left\{-\frac{1}{2}(x+y)\right\} \qquad x,y > 0$$

and zero otherwise, for some normalizing constant c_1.

Consider continuous random variable U defined by

$$U = \frac{1}{2}(X - Y).$$

Find the pdf of U, f_U.

2. In biology, a (2-D) confocal microscopy image of a cell nucleus is well represented by an ellipse with parameters $a > b$. Within the cell nucleus are found localized protein bodies (called PMLs), and a key biological question relates to the spatial distribution of the PMLs in the nucleus.

Suppose that the (x, y) coordinates of a PML body in the image of a nucleus (suitably rotated and standardized for magnitude) are continuous random variables X and Y with joint pdf

$$f_{X,Y}(x,y) = c_2 \qquad \frac{x^2}{a^2} + \frac{y^2}{b^2} < 1$$

and zero otherwise for some normalizing constant c_2 (that is, the pdf is constant on interior of the ellipse, and zero otherwise).

(a) Find the marginal pdfs for X and Y implied by this joint model.

(b) Show that the covariance between X and Y is zero, and hence that the two variables are uncorrelated.

(c) Are X and Y independent ? Justify your answer.

3. (a) Compute, from first principles, the correlation

$$Corr_{f_{X,Y}}[X,Y]$$

when

$$X \sim Normal(0,1)$$

and $Y = X^2$.

Are X and Y independent ? Justify your answer.

Hint: If $Y = X^2$, the rules of expectation dictate that for a general function h

$$E_{f_{X,Y}}[h(X,Y)] \equiv E_{f_X}\left[h\left(X,X^2\right)\right]$$

(b) Suppose that X_1 and X_2 are independent standard normal random variables. Define random variables Y_1 and Y_2 by the multivariate linear transformation

$$Y = AX + b$$

where $X = (X_1, X_2)^T$ and $Y = (Y_1, Y_2)^T$ are the column vector random variables, A is the 2×2 matrix

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$$

and $b = (1,2)^T$ is a constant column vector.

(i) The marginal distribution of Y_1.
(ii) The covariance and correlation between Y_1 and Y_2.

4. (a) Suppose that X_1 and X_2 are independent and identically distributed continuous random variables with cumulative distribution function

$$F_X(x) = \frac{x}{1+x} \qquad x > 0$$

with $F_X(x) = 0$ for $x \leq 0$.

Find $P[X_1 X_2 < 1]$.

(b) Suppose that Z_1 and Z_2 are independent $Normal(0,1)$ random variables. Let

$$Y_1 = \frac{Z_1}{\sqrt{Z_1^2 + Z_2^2}} \qquad Y_2 = \sqrt{Z_1^2 + Z_2^2}.$$

Find the marginal probability density function of Y_1.

Are Y_1 and Y_2 independent ? Justify your answer.

5. (a) Show how an exponential family distribution can be constructed by *tilting* a pdf f_X.

(b) Let f_Y be a pdf for random variable Y, and let $s(Y)$ be a transformed version of Y such that $Var_{f_S}[s(Y)] > 0$. Let the set $\mathcal{N}$ be defined by

$$\mathcal{N} \equiv \left\{ \theta \in \mathbb{R} : K_S(\theta) = \log \left[\int e^{s(y)\theta} f_Y(y)\, dy \right] < \infty \right\}$$

(i) Show that $0 \in \mathcal{N}$.

(ii) Using Hölder's Inequality, show that $\mathcal{N}$ is a **convex set**, that is, if $0 \leq \alpha \leq 1$ and $\theta_1, \theta_2 \in \mathcal{N}$, then

$$\alpha\theta_1 + (1-\alpha)\theta_2 \in \mathcal{N}.$$

(iii) Show that $K_S(\theta)$ is a **convex function** on $\mathcal{N}$, that is, if $0 \leq \alpha \leq 1$ and $\theta_1, \theta_2 \in \mathcal{N}$, then

$$K_S(\alpha\theta_1 + (1-\alpha)\theta_2) \leq \alpha K_S(\theta_1) + (1-\alpha) K_S(\theta_2)$$

6. (a) Suppose $X_1, \ldots, X_n, \ldots$ are a sequence of random variables with cumulative distribution functions defined by

$$F_{X_n}(x) = \left(\frac{1}{1+e^{-x}} \right)^n \qquad x \in \mathbb{R}.$$

Find the limiting distributions as $n \longrightarrow \infty$ (if they exist) of the random variables

(i) X_n,

(ii) $U_n = X_n - \log n$.

Using the result in (ii), find an approximation to the probability

$$P[X_n > k]$$

for large n.

(b) In a dice rolling game, a fair die (with all six scores having equal probability) is rolled repeatedly and independently under identical conditions. On each roll, the player wins six points if the score is a 6, loses one point if the score is either 2,3,4 or 5, and loses two points if the score is 1.

Let T_n denote the points total obtained after n rolls of the die. The player begins the game with a points total equal to zero, that is $T_0 = 0$.

(i) Find the expectation and variance of the points total after 100 rolls of the die.

(ii) Find an approximation to the distribution of the points total after n rolls, for large n.

(iii) Describe the behaviour of the sample average points total, $M_n = T_n/n$, as $n \longrightarrow \infty$.

7. (a) (i) Suppose that random variable X has a Poisson distribution with parameter λ. Show that standardized random variable,

$$Z_\lambda = \frac{X-\lambda}{\sqrt{\lambda}} \xrightarrow{d} Z \sim N(0,1)$$

as $\lambda \rightarrow \infty$.

(ii) Suppose that $X_1, ...X_n \sim Poisson\,(\lambda_X)$ and $Y_1, ...Y_n \sim Poisson\,(\lambda_Y)$, with all variables mutually independent. Find μ such that the random variable M defined by

$$M = \overline{X} + \overline{Y}$$

satisfies

$$M \xrightarrow{p} \mu$$

where

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

are the sample mean random variables for the two samples respectively.

(b) Suppose that $X_1, ..., X_n \sim Exponential(\lambda)$. The cdf of the random variable $T_n = \max\{X_1, ..., X_n\}$ is given by

$$F_{T_n}(t) = \{F_X(t)\}^n .$$

where F_X is the cdf of $X_1, ..., X_n$.

(i) Find $F_{T_n}(t)$ explicitly.

(ii) Discuss the form of the limiting distribution of T_n as $n \longrightarrow \infty$.

(iii) Find the form of the limiting distribution of random variable U_n, defined by

$$U_n = \lambda T_n - \log n$$

as $n \longrightarrow \infty$.

8. Suppose that $X_1, X_2, ...$ are i.i.d $Cauchy$ random variables with pdf

$$f_X(x) = \frac{1}{\pi}\frac{1}{1+x^2} \qquad x \in \mathbb{R}$$

and characteristic function $C_X(t) = \exp\{-|t|\}$.

(a) Find the distribution of the random variable T_n defined by

$$T_n = \sum_{i=1}^{n} X_i.$$

(b) Describe the behaviour of the sample mean statistic

$$\overline{X}_n = \frac{T_n}{n}$$

as $n \longrightarrow \infty$, quoting any theorems that you rely on.

(c) Show that the Cauchy distribution can be constructed as a *scale mixture* of a normal distribution with a Gamma mixing distribution.